

**UNITED NATIONS STATISTICAL COMMISSION and
ECONOMIC COMMISSION FOR EUROPE**

CONFERENCE OF EUROPEAN STATISTICIANS

Work Session on Statistical Data Editing

(Vienna, Austria, 21-23 April 2008)

Topic (iii): Improvement of quality through data editing

EDITING PROJECT OF STATISTICS FINLAND

Supporting Paper

Prepared by Pauli Ollila, Statistics Finland

Abstract

Editing and imputation practices at Statistics Finland have been developed during the years mainly on the practical basis, varying from one statistics to another. Recently there has been a growing need for more information about general principles and good practices of editing and its quality evaluation. The article describes the main aspects of the editing project planned to be started in 2008. The objectives of the project are to provide an applicable model for the editing process including indicators for quality, software tools for that purpose and both theoretical and practical education on editing and imputation. When ready, the results of this project should help the researchers with their editing problems, so a close collaboration with the target groups is essential during the project. Furthermore, one of the key factors in this project is to study the development work of this area at some statistical offices and to tie contacts with these organisations. The benefits of the project will be the improvement in the quality of statistics together with systematic processing saving time and resources allocated to the editing work.

Keywords: editing, imputation, quality

I. BACKGROUND

1. One of the objectives of Statistics Finland is the intention to have the production of statistics at the best possible level, and one of the production phases for this purpose is editing of the data material collected for statistics. The editing practices vary from one statistics to another. Often they are developed to meet the purposes of the statistics during a long time, and these actions are included to different phases when needed. The development of information technology has eased and speeded up editing remarkably. On the other hand, the quality of the editing at Statistics Finland cannot be sufficiently evaluated, because there are no clear measures for that purpose. Recently there has been a growing need in different branches of statistics to get more knowledge about the general principles of editing and good editing practices. On the other hand there has been some concern about the fact that with current practices editing might use the resources of the statistical office too much. There should be methods and practices which can be used for improving the quality of the statistical information in the production process more efficiently.

2. There have been plans to start an editing project at Statistics Finland in 2008. This paper outlines some main aspects of these plans. It must be noted that this project is not yet set and the ideas behind it might still be revised.

II. MAIN ASPECTS

3. If the project is realised, its aim is to create a model for editing process, suitable for the methods of Statistics Finland, and enabling the unification and clarification of editing practices and also the evaluation of the quality of editing with concrete measures. Furthermore, the aim is to outline the operations for the implementation, to plan the education of editing and create a framework for the realisation of the education. The target group of the project work is those branches at Statistics Finland, where editing is carried out.

4. The objectives of the project could be dealt into theoretical definition work, preparing application work and education. In order to set clear ways of action and quality measures, there must be some kind of model for editing process. This requires some study work about the existing international models, one starting point e.g. the process description of editing in New Zealand. At the moment, editing is often conducted with a principle "all at once", when it is assumed that editing is taking place when running one specific program and after that everything is all right. The goal is to present editing as a process, which will be carried out one part at a time and after each part there are checks whether the material is ok.

5. The model of the editing process includes some study criteria and editing practices. Especially when considering the quality evaluation of editing there should be measures of quality which have a concrete interpretation. The goal is to include the information about the quality of editing and other quality measures as a part of the process. The project should do this development work considering the practical situation in different statistics. Similarly there will be a connection to other quality activity at Statistics Finland: the quality measures and especially the practices of editing and information on their quality should be included in the metadata consisting of all definition of statistical information.

6. The target of the project is to define prerequisites and ways of action in the beginning of applying and developing the necessary program tools for editing and imputation. For editing purposes program tools will be purchased, with which editing can be conducted by using graphical interfaces and program-like components applicable for production. One example of this is the SAS-based software BANFF from Statistics Canada. With these tools we support the development of editing as a process, which improves the quality of the statistics and this quality can be immediately evaluated. Special attention is paid to the fact that the software solutions produce clear information which measures the success and quality of editing. The adoption of the editing process as well as the editing methods and software requires education targeted to the personnel carrying out editing in different statistics of Statistics Finland. The purpose is to prepare the contents of this education and to create the framework for this education.

7. The benefits of the project will be the improvement of the quality of making statistics as a consequence of a systematic checking and correction process, time savings following of the removal of some slow and partially manual phases of activities and the cost reductions due to the diminishing resources needed for editing (possibility to move those resources to other phases). The end result of the project should be concrete practices for editing, tool(s) for realising the practices and measures and information on the implementation.

III. COOPERATION

8. The developmental work of editing processes, studies on software applications and adjusting the methodological know-how require collaboration with several co-operational groups. At the international level it is essential to tie connections with statistical offices which are conducting / have conducted development projects for editing or related issues (e.g. Sweden, Canada and New Zealand), and to be active in the international groups with scientific and practical aims, e.g. the work sessions of statistical editing within UNECE. At the national level there are some active persons at the universities concerning the editing and imputation issues. The target groups and the statistical branches implementing editing at Statistics Finland should be consulted frequently in order to ensure the applicability. It is natural that some members of the methodological unit of Finland and the SAS experts are in co-operation with the project. Also the

development projects related to the issue are good for contact especially when finding out the needs of the statistics concerning editing.

IV. STRUCTURE

9. The timetable of the project is planned to cover two years. However, the duration of the project will be evaluated in the first "checkpoint" of the project, about a year after the beginning. In that point the first results are evaluated together with the prerequisites of the project, e.g. concerning the allocation of the work time available in the project group. It is crucial that the preliminary studies of the situation and needs of editing are begun straight at the beginning of the project.

10. Connections to possible existing development projects are tied. Furthermore, the collaboration between different co-operational groups should be started as soon as possible, both at the national and international levels. The developmental work will have gradually increasing role when the preliminary studies have been done and the connections to the co-operational groups are well established. One of the first aims is to create a process flow chart applicable for Statistics Finland and to write the descriptions it would include. After assessing the framework in which the process would take place the next step is to specify the requirements for the software solution to be used. The results concerning the methodology and the practices together with the software applications are presented in the education phase, though some consulting of the methods can be conducted earlier as well. The project does not include exact programming work of the software applications needed for editing.

11. The project group of size 6 to 8 is planned to include representants from the methodological unit, the unit of information technology and some statistics. The last connection is especially important in order to keep the project targeted to the real problems of the practicing statisticians. The management group would include one director from some statistical branch, one methodologist possibly from the university and one high-level specialist on information technology. The support group would consist of various experts on SAS, editing procedures (e.g. BANFF), persons practicing editing and experts on statistical methods.